



# Solving Power Design Challenges in the Age of AI

Jeff Chang

# What's happened – 2016 to 2024

Process node  
16 to 4nm

16X

TDP

300 to 1000W

3.3X

Transistors  
15.3 to 208B

13.6X

TDC

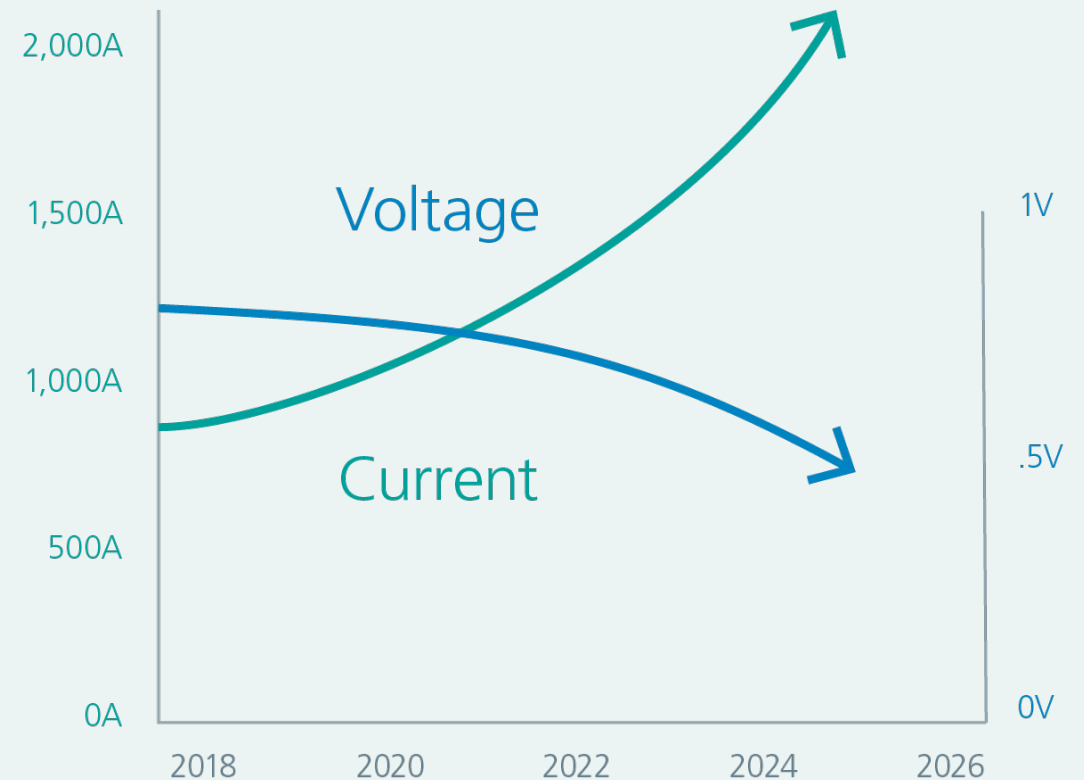
200 to 1000A ↗

5X

# Powering high performance processors

- Modern processors (GPU, CPU, NPU) need a lot of current...
- Increasing PDN distribution losses  
Decreases in power efficiency
- Significant operating performance reduction if power demands are not met
- Added complexity of decreasing operating voltages as move continue to lower fabrication nodes

Progression of processor peak current requirements and lower operating voltages



# Conduction Loss in PCB

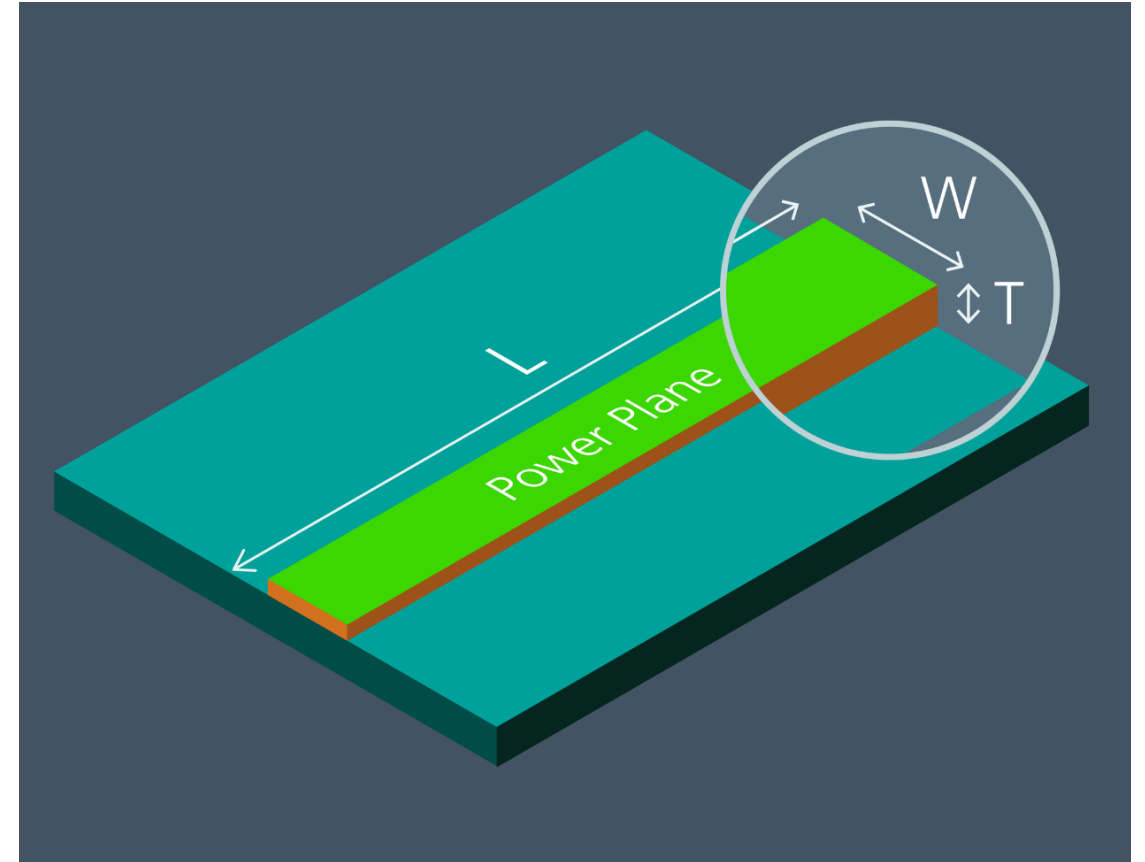
## ■ PCB Trace Resistance Equation:

$$R = \frac{\rho L}{TW} [1 + \alpha(T_{amb} - 25^{\circ}C)]$$

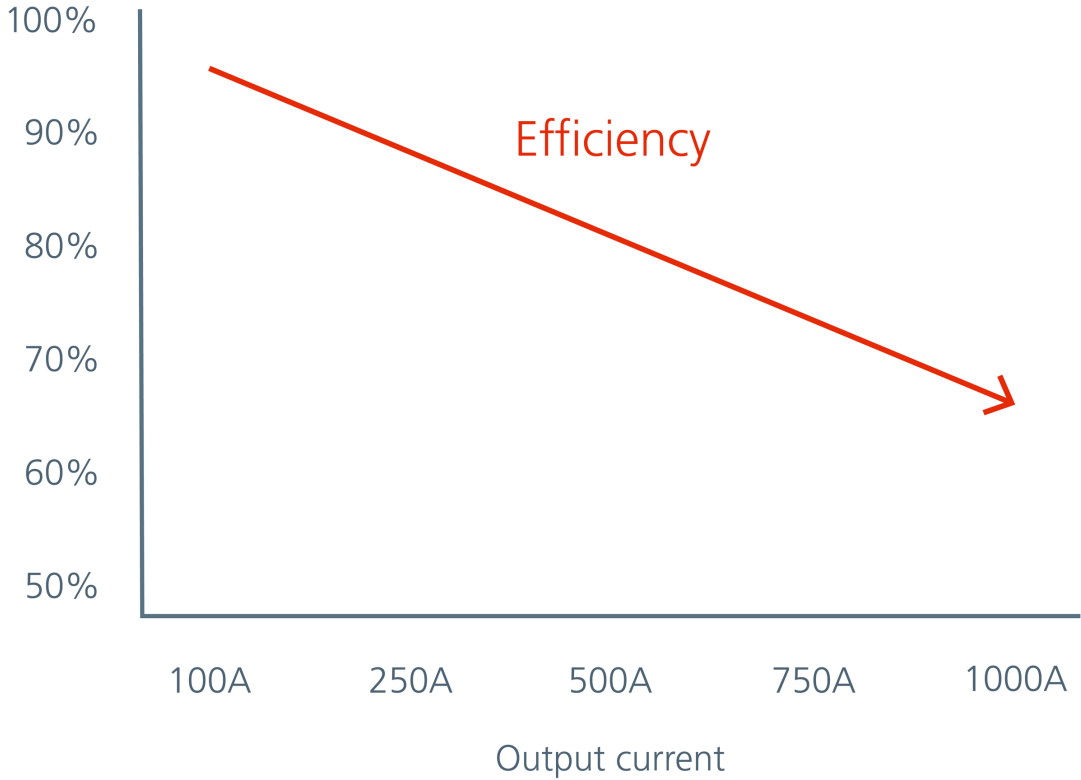
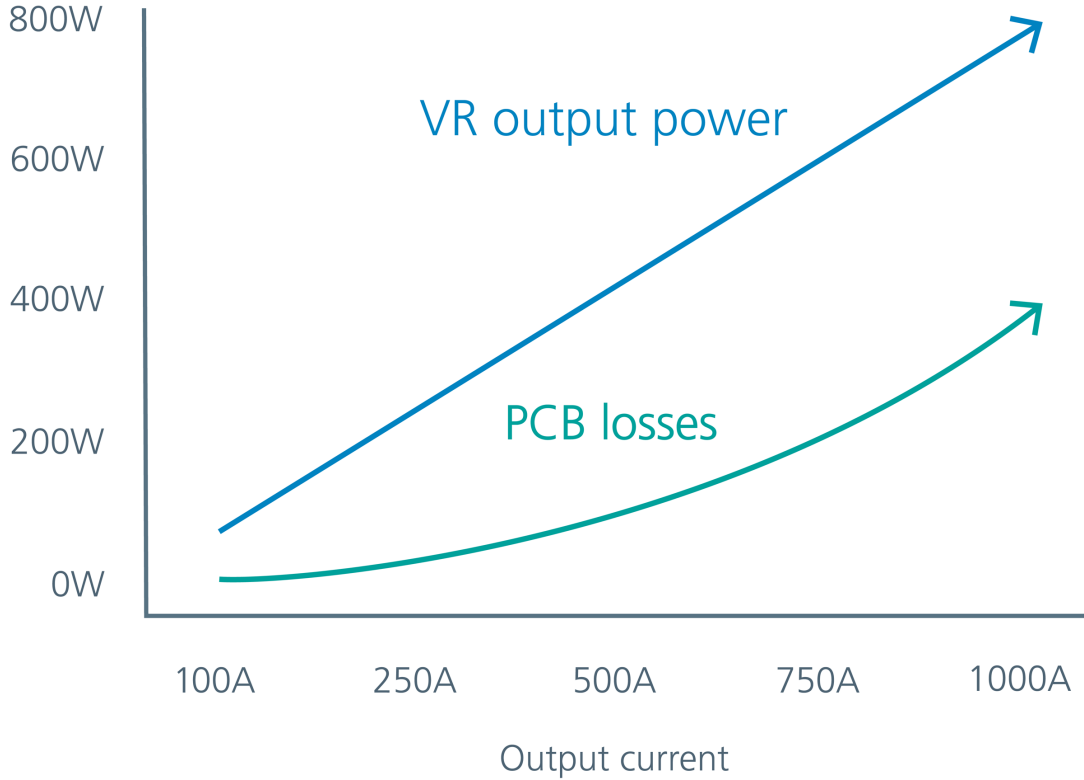
- R is the resistance
- L is the length of the trace
- T is the thickness of the trace
- W is the width of the trace
- Tamb is the ambient temperature

## ■ Ohm's Law:

$$R = \frac{V}{I}$$

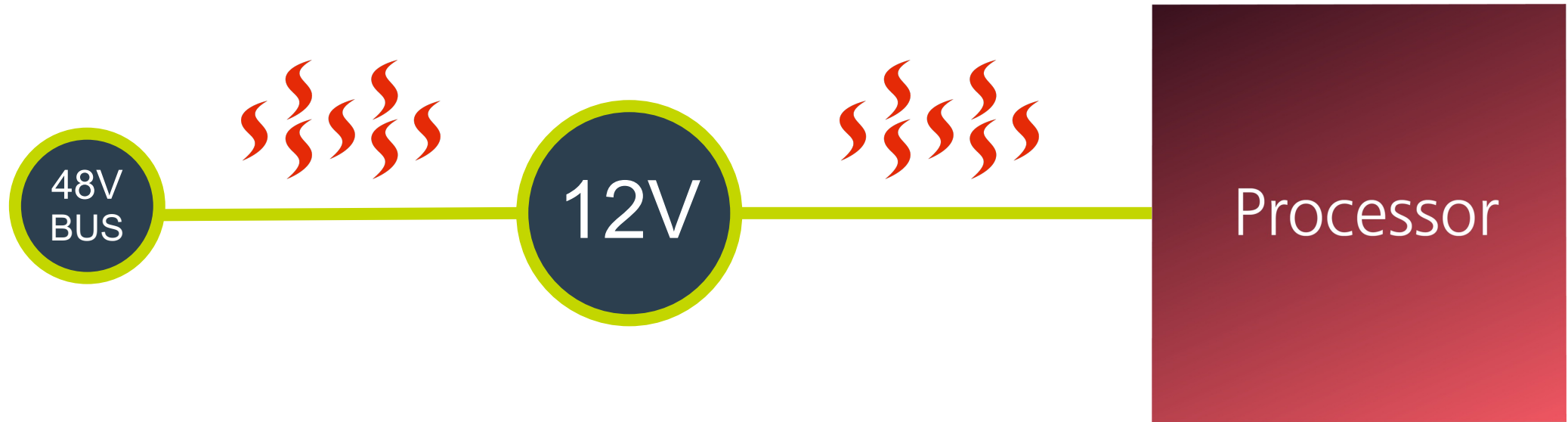


# Voltage Regulator to the Processor Losses



Example with PCB resistance of 400uOhm (VR at 0.8Vout)

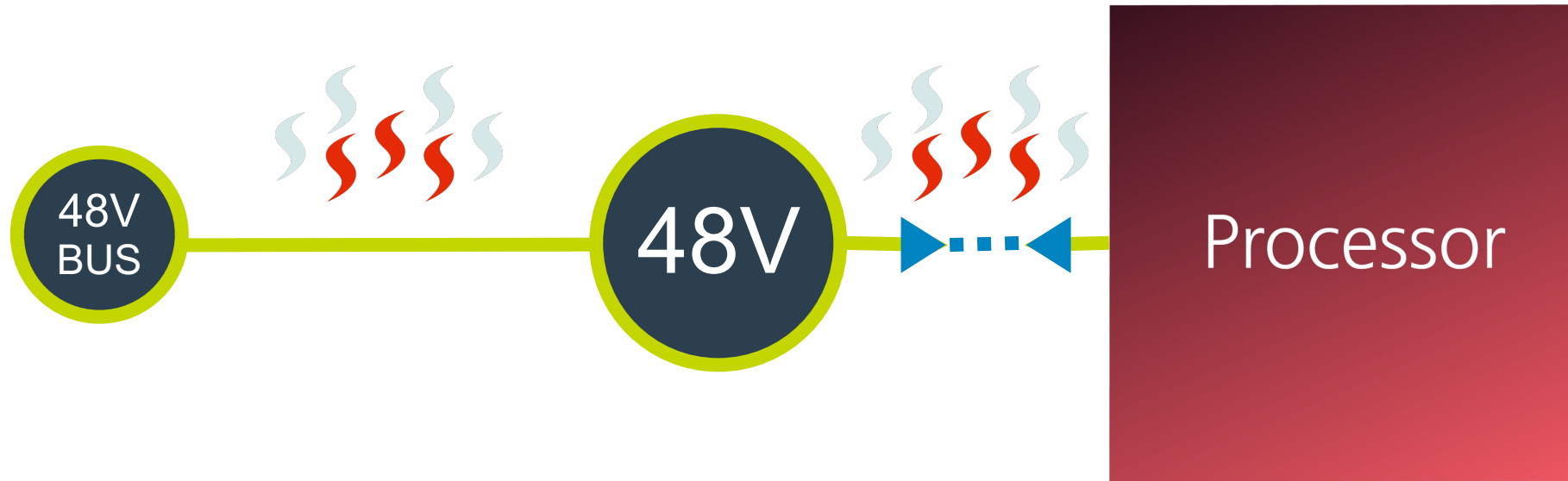
# Power Delivery with Different Structure



# Power Delivery with Different Structure

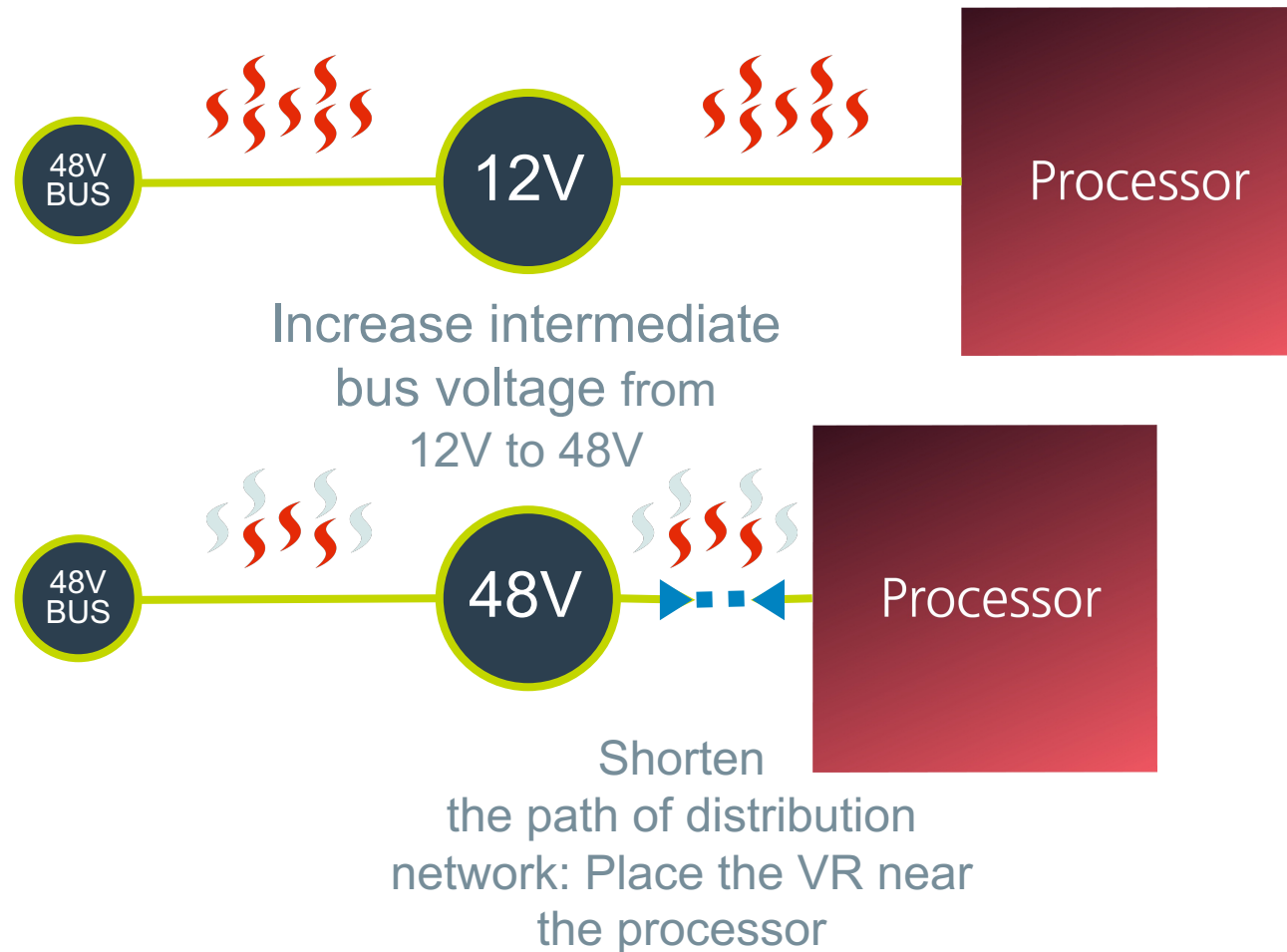


# Power Delivery with Different Structure



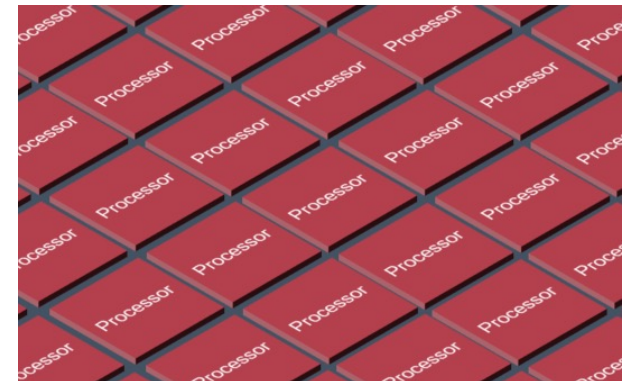
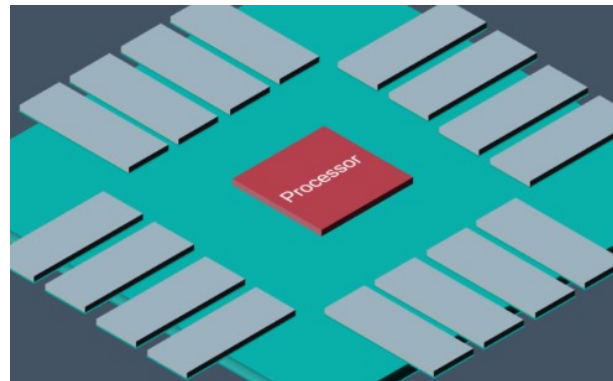
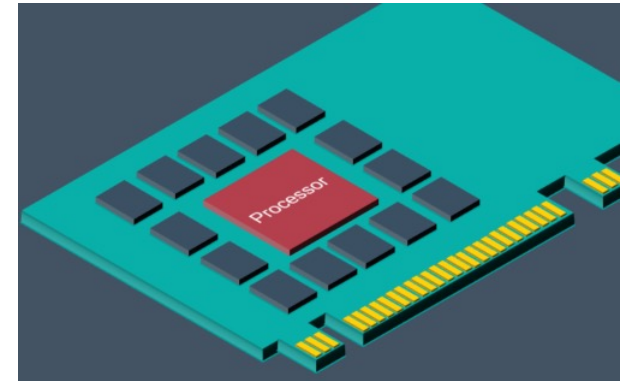
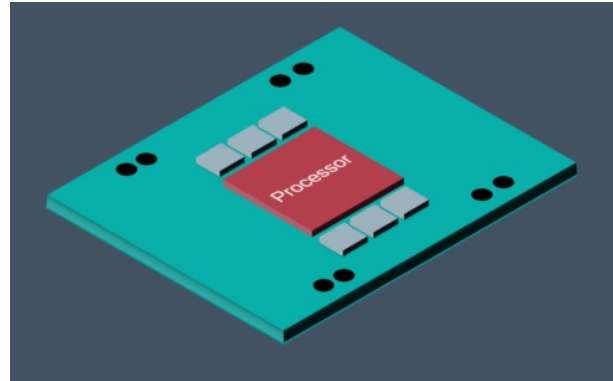


# Power Delivery with Different Structure



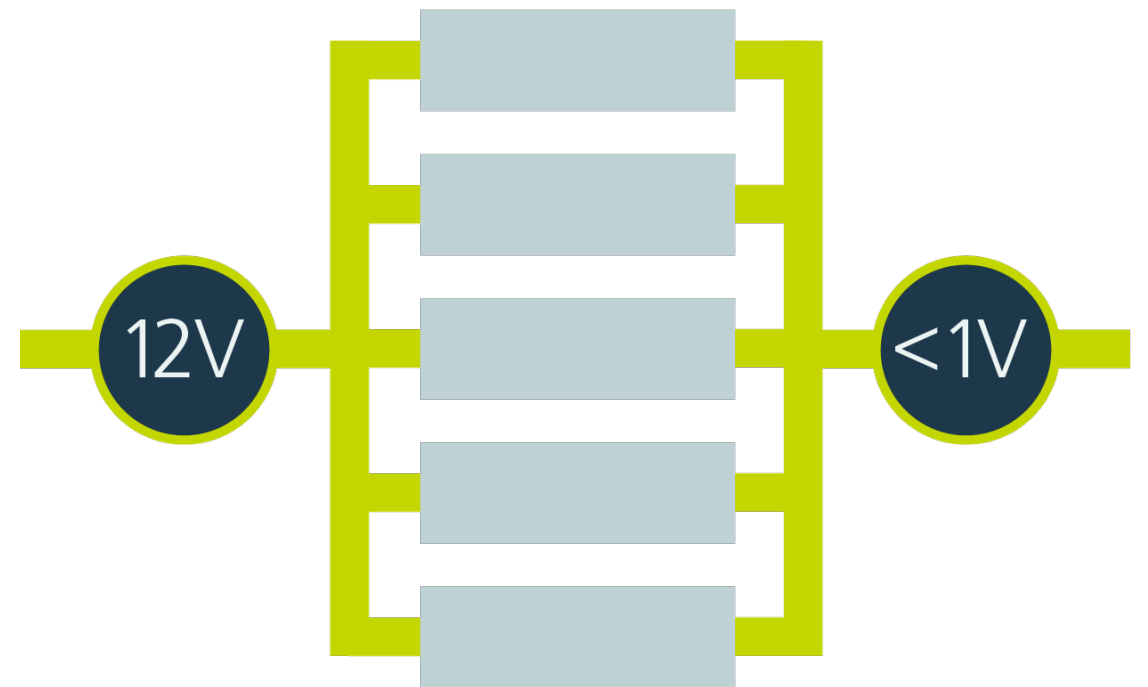
# Space Constraints Challenge

- OAM and custom AI accelerator cards
  - PCB size limitations
  - On-board memory blockage
- Network switch processors
  - High speed transceiver blockage
- Cluster computing
  - Wafer Scale Engine
  - PCB level xPU grid fabric



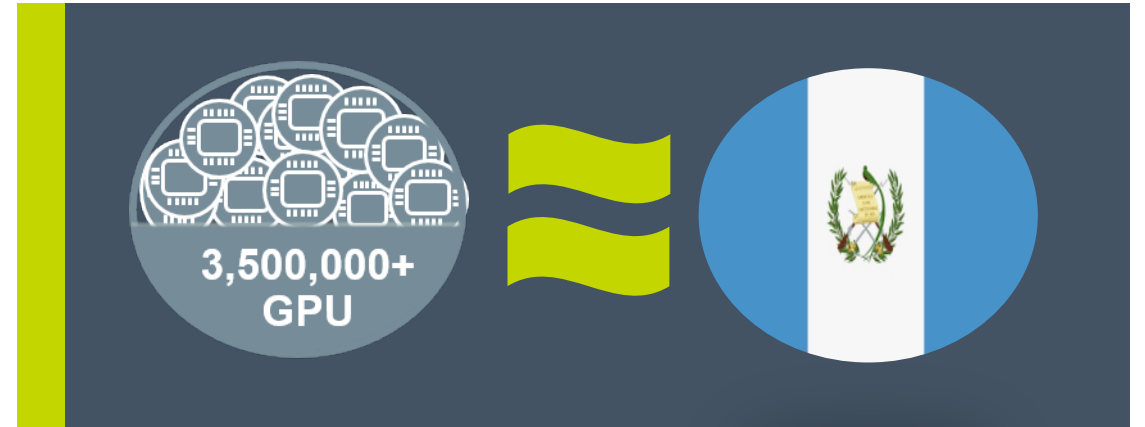
# Conventional multiphase

- Conversion performed by DrMOS/Inductor
- High conversion ratio (min. 12:1)
- Challenging to scale for higher currents
- Phase unbalancing
- Noise generation
- Size prohibits reducing PDN



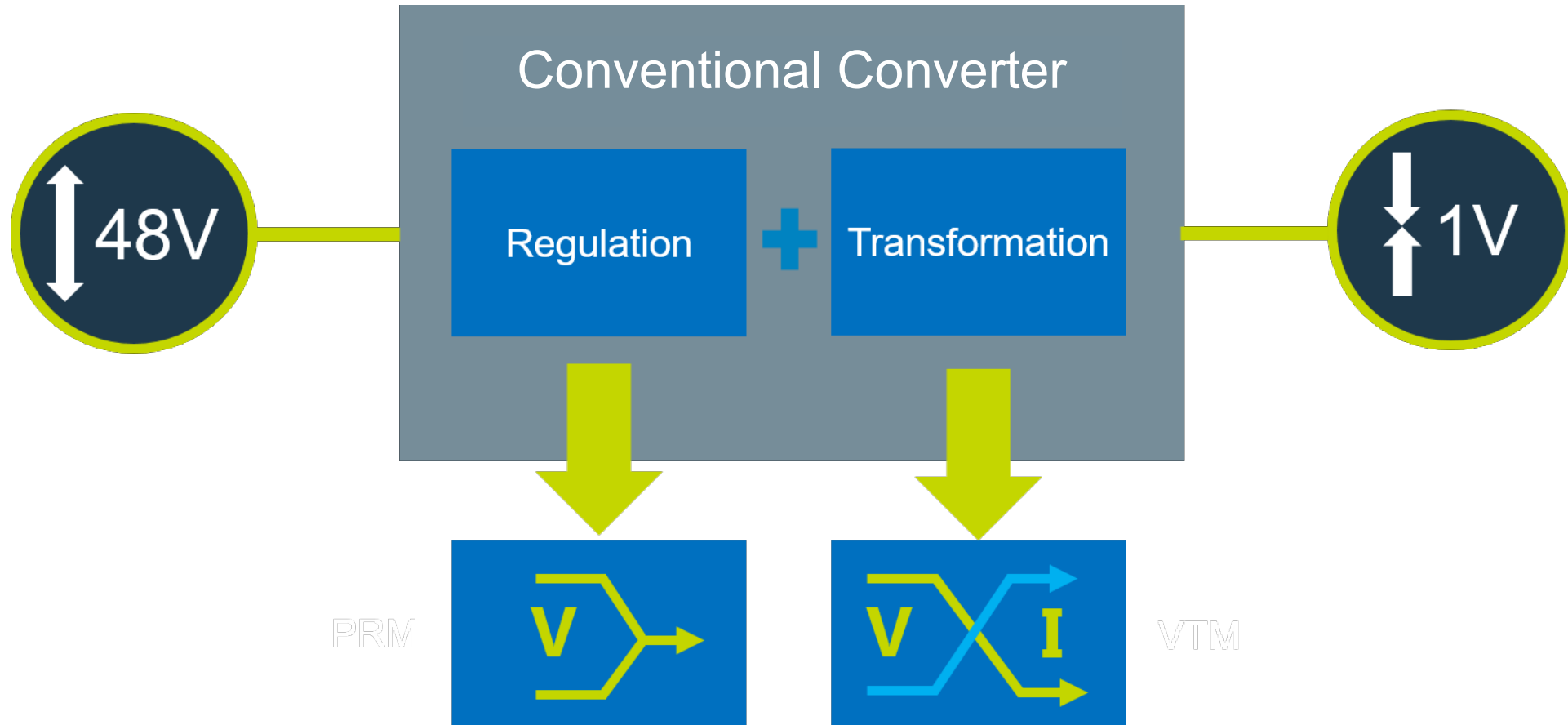
# Global Energy Crisis

# Global Energy Crisis

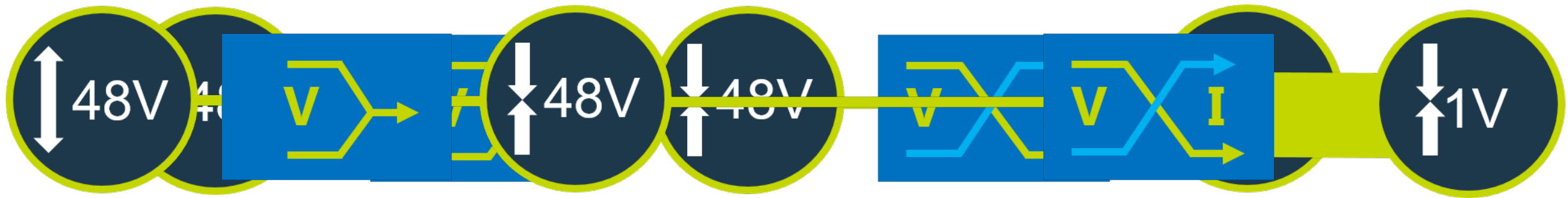


# Factorized Power Architecture

# Factorized Power Architecture



# Factorized Power Architecture



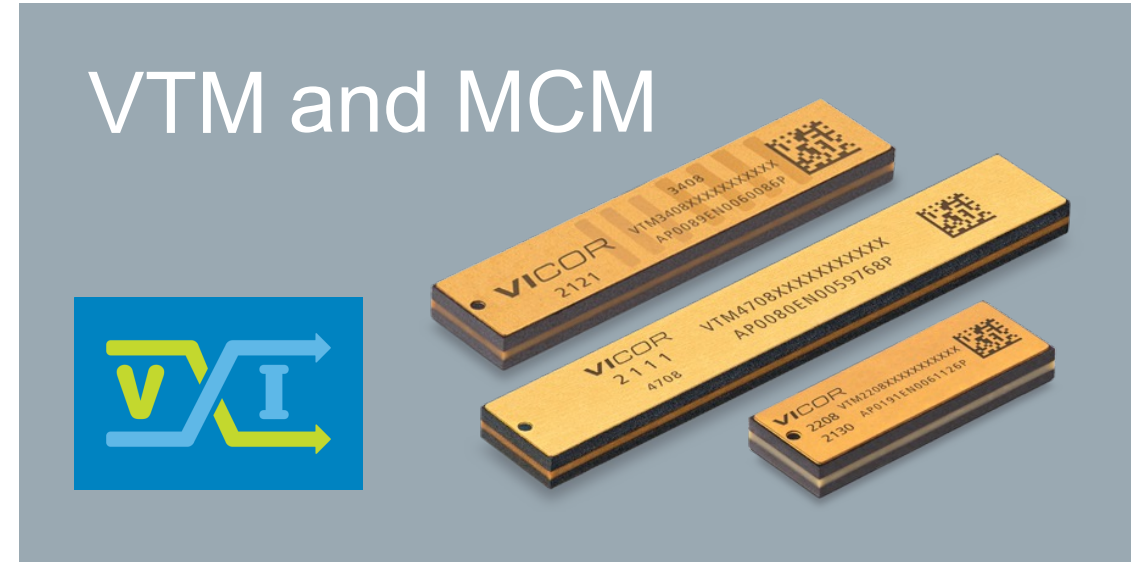


# Flexible Product Specifications for Versatile Combinations



### Variety of PRMs

- 20x10mm = 250W
- 23x14mm = 500W
- 37x18mm = 1000W



### Scalable VTMs

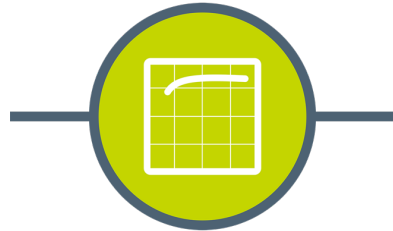
- 22x8mm = 125A
- 34x8mm = 250A
- 42x8mm = 325A

## NEW

### Scalable MCMs

- 100A~ 1000A

# Factorized Power Architecture



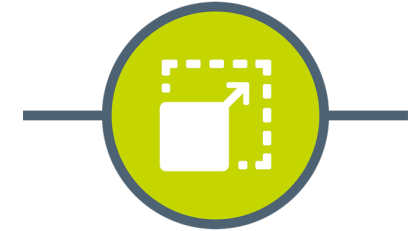
Minimizes  
wasted energy



Low EMI



Adjusts for any load



Scales for power  
demands



High density



Component position  
flexibility



Fast transient response

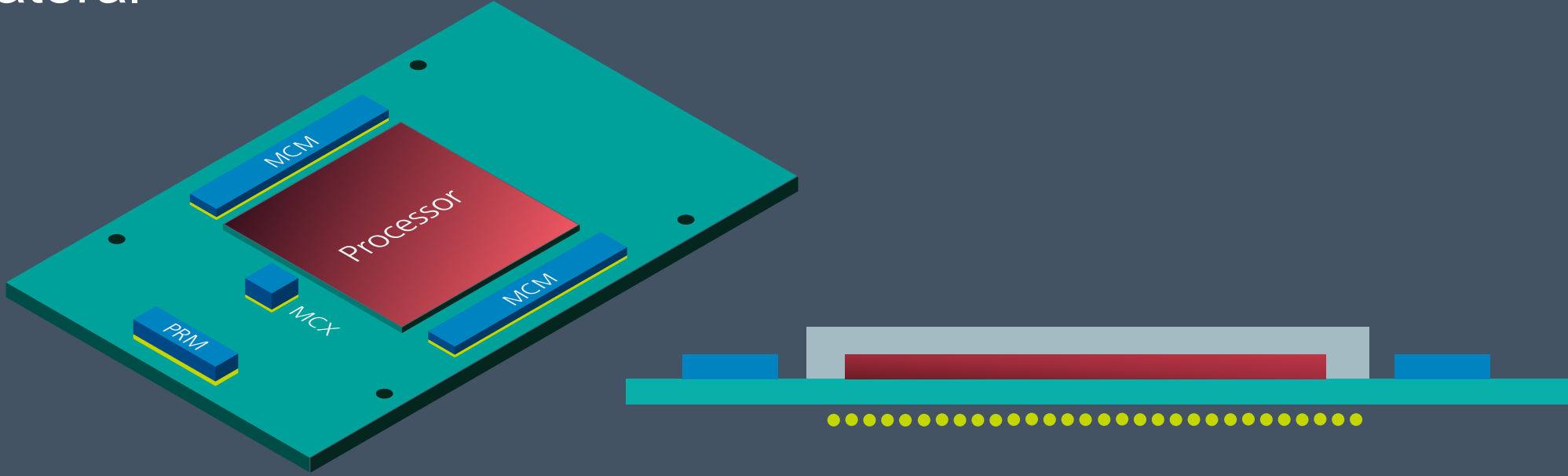


Easy, modular design

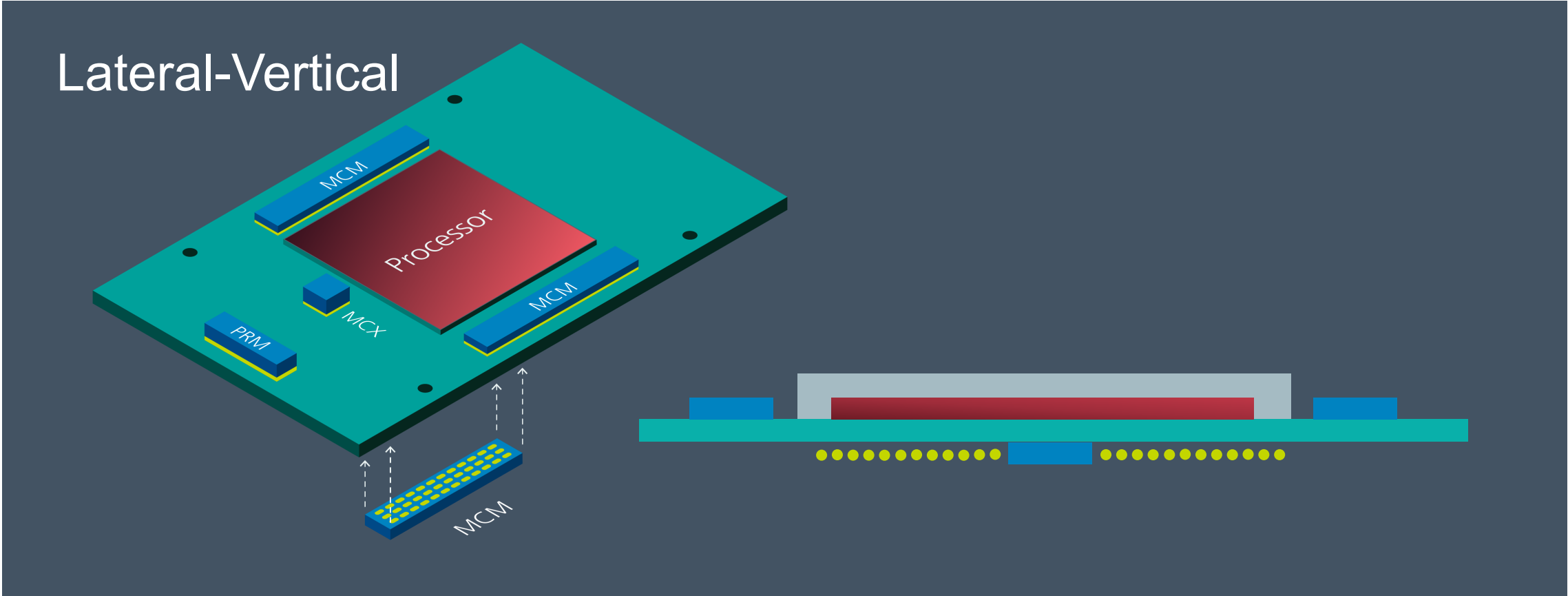
# Power Delivery Networks

# With Different Power Delivery Architecture

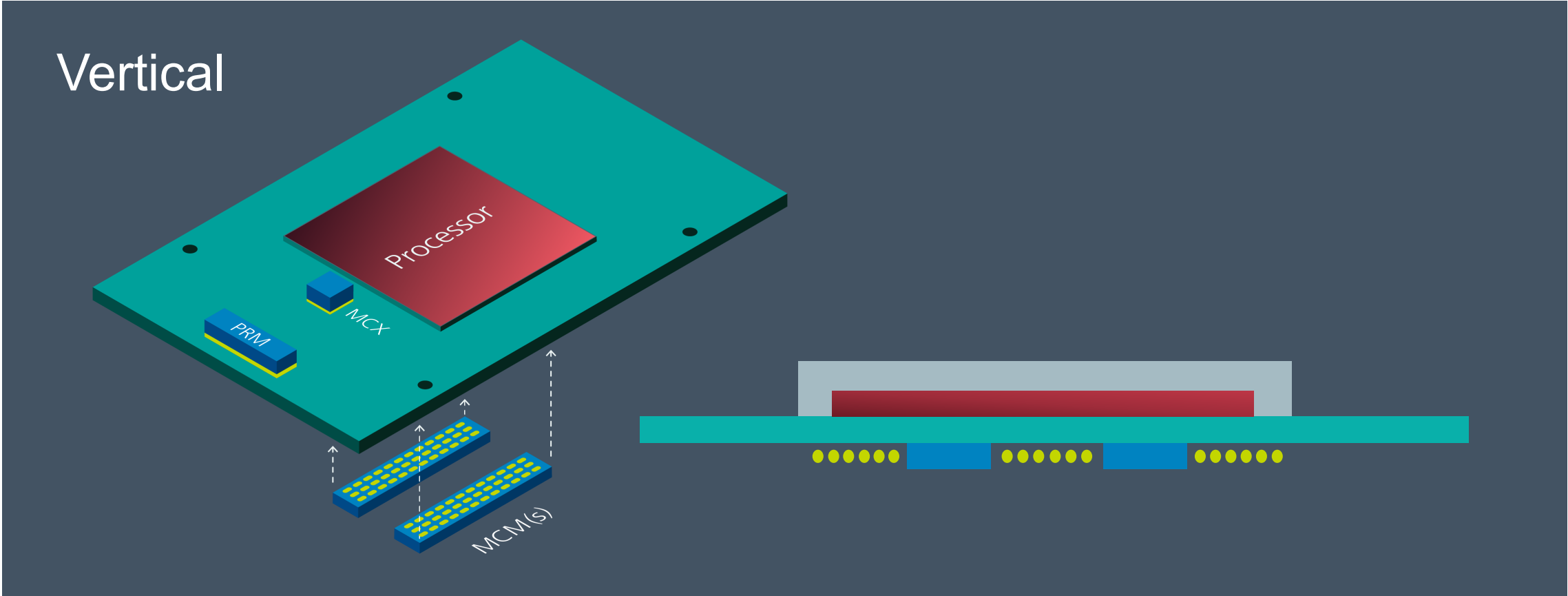
Lateral



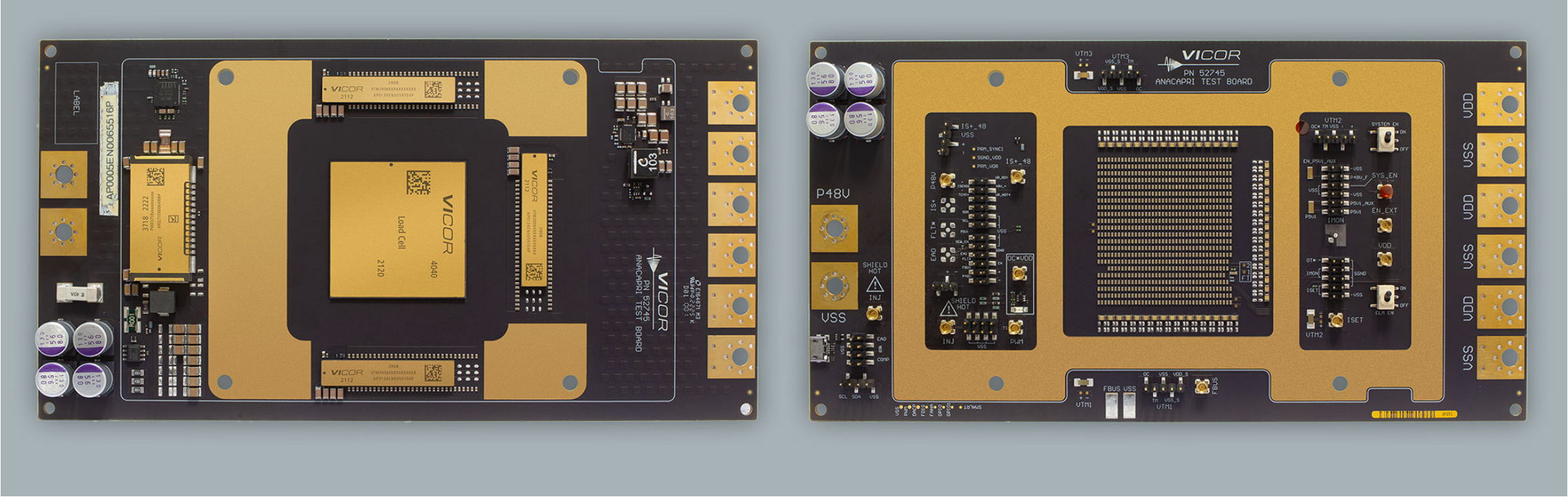
# With Different Power Delivery Architecture



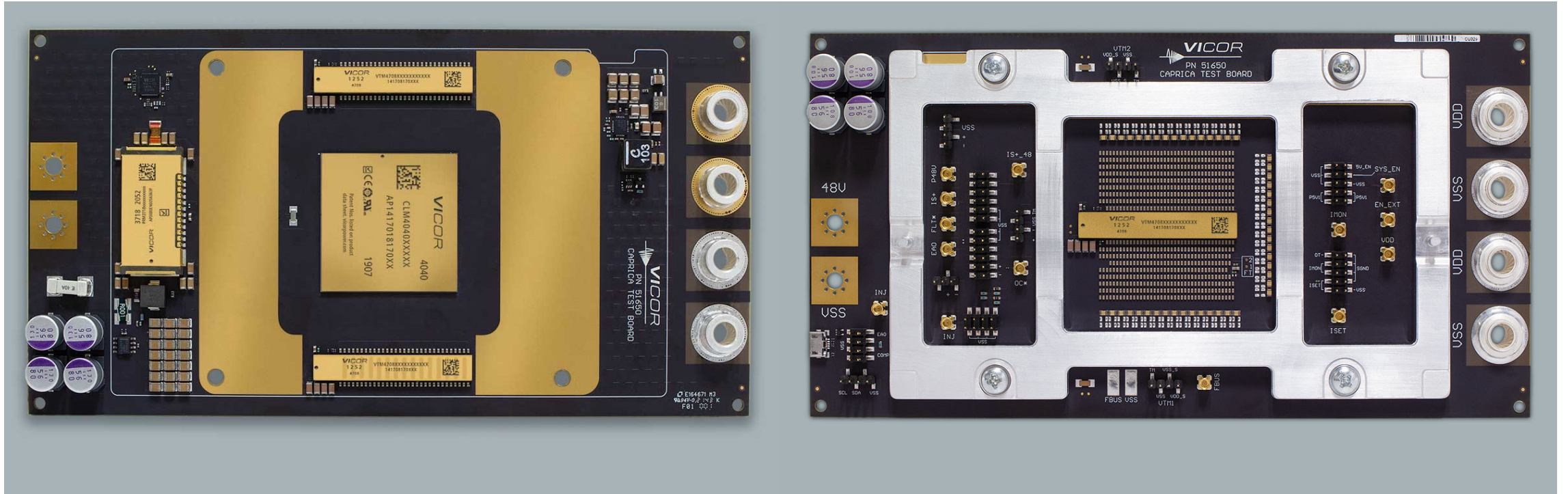
# With Different Power Delivery Architecture



# Lateral Power Delivery



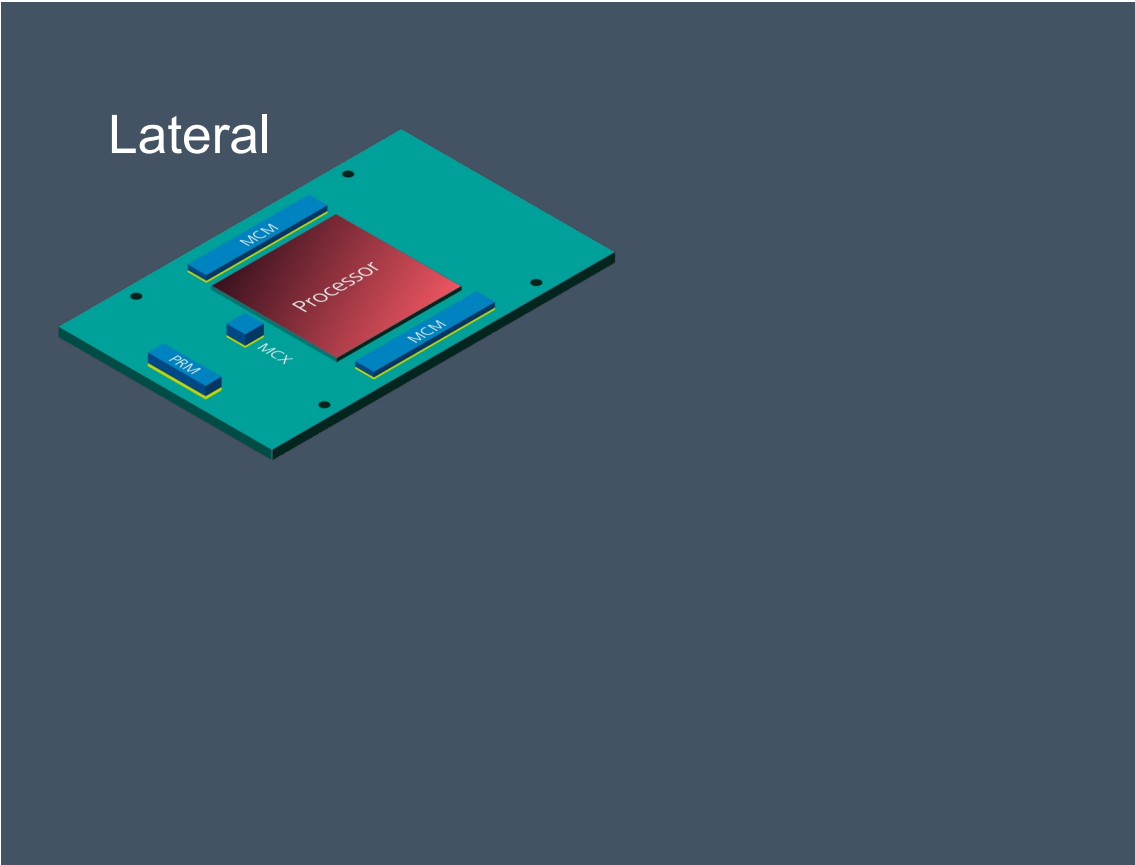
# Lateral-Vertical Power Delivery



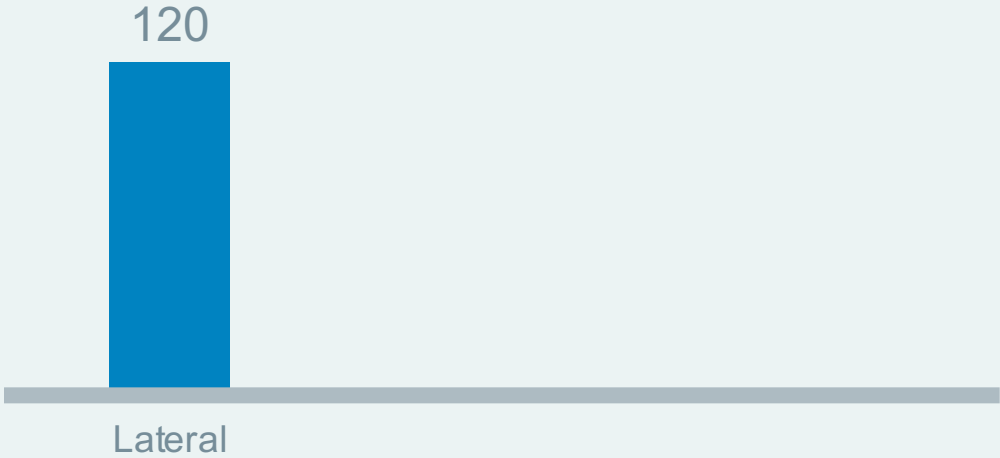


# Performance loss analysis

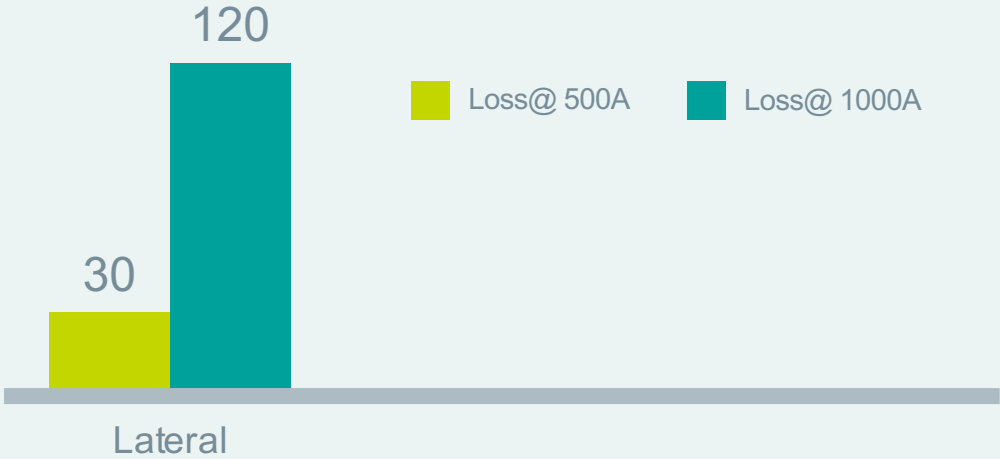
# Performance loss analysis



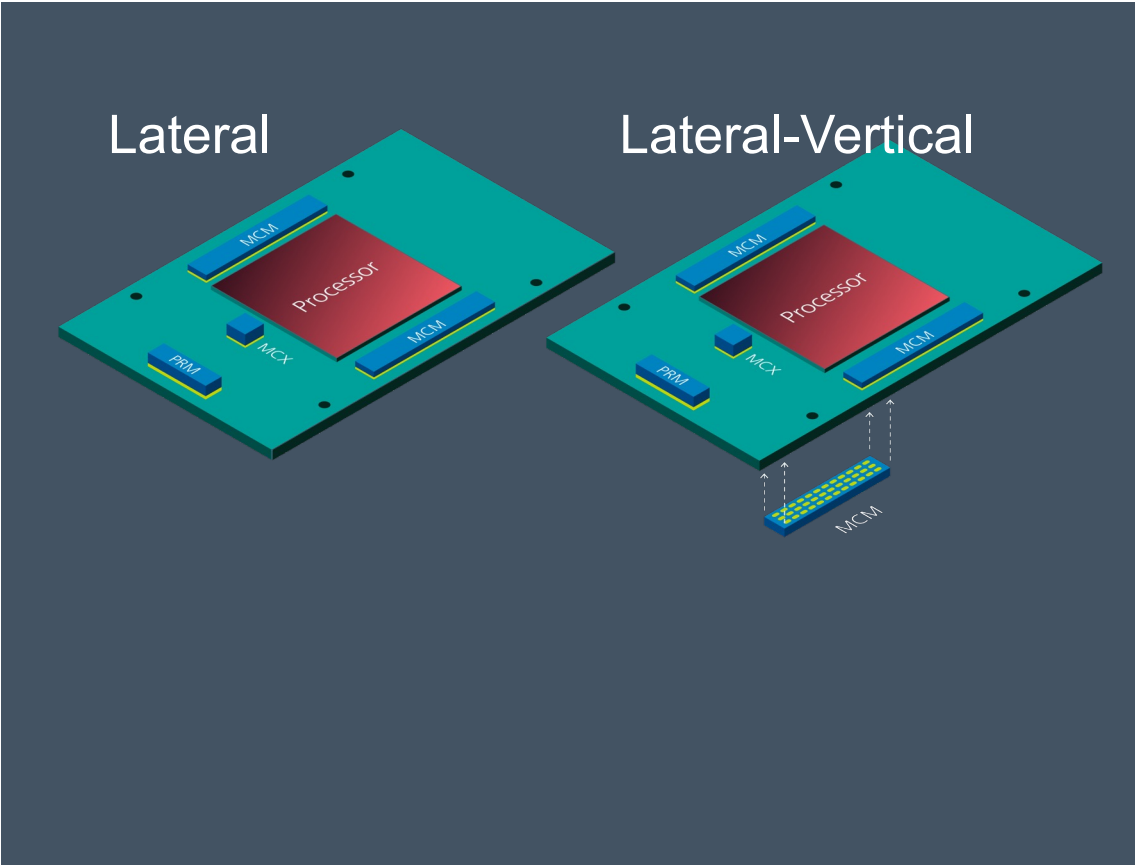
PDN Resistance (Ohms)



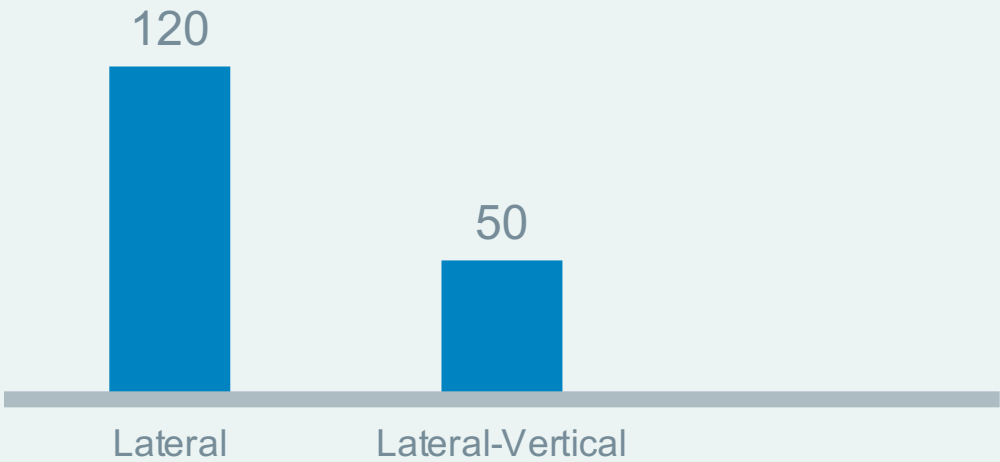
PDN Loss (Watts)



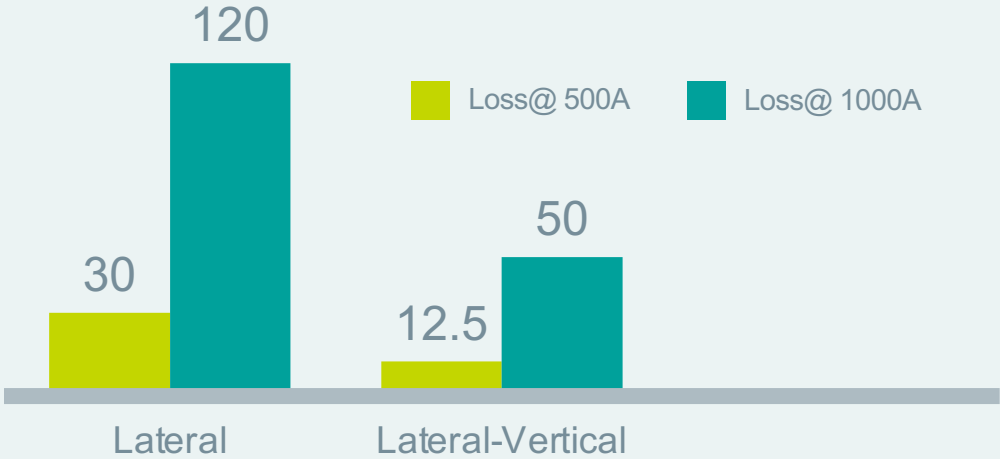
# Performance loss analysis



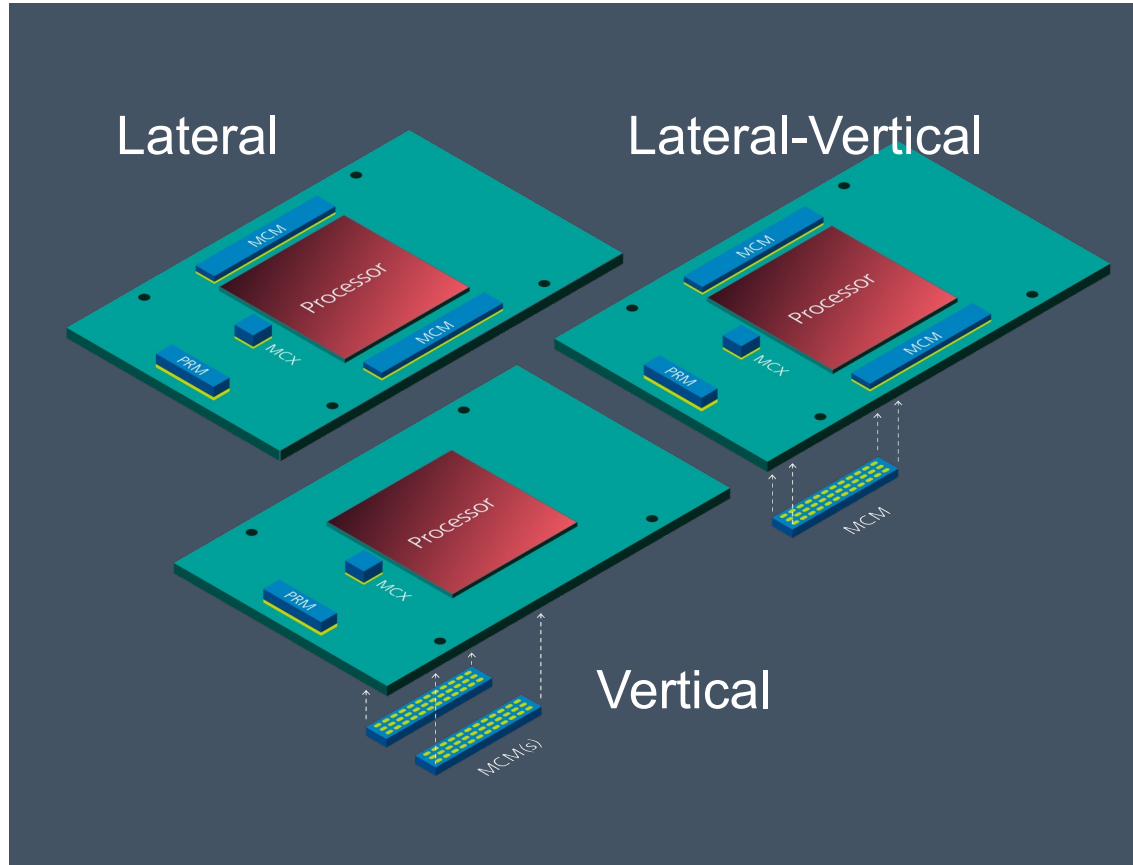
PDN Resistance (Ohms)



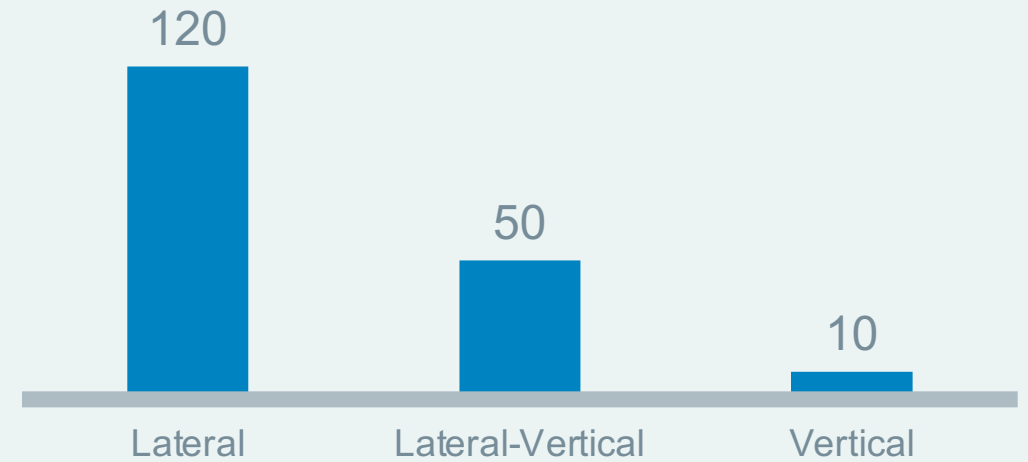
PDN Loss (Watts)



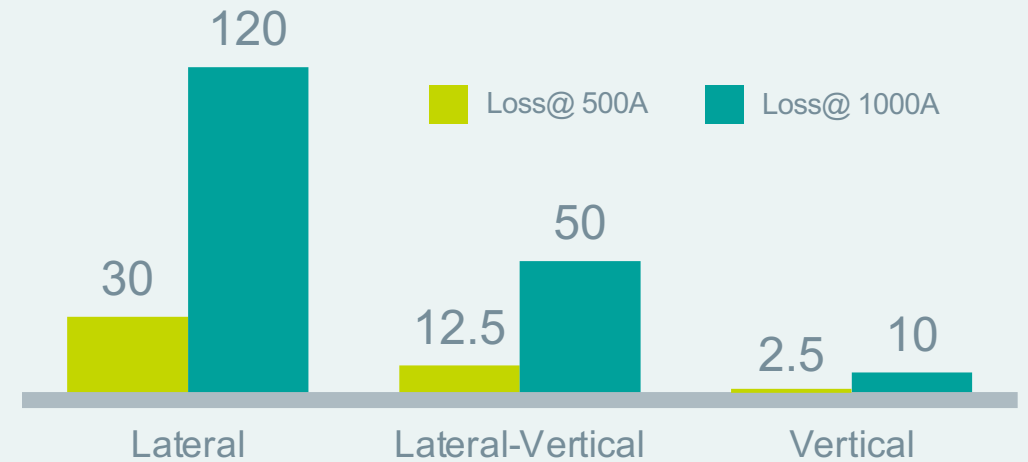
# Performance loss analysis



PDN Resistance (Ohms)



PDN Loss (Watts)



*VICOR*

Thank you

# Reference

- <https://www.tomshardware.com/tech-industry/nvidias-h100-gpus-will-consume-more-power-than-some-countries-each-gpu-consumes-700w-of-power-35-million-are-expected-to-be-sold-in-the-coming-year>
- <https://countryeconomy.com/energy-and-environment/electricity-consumption>